A PROOF OF THE INVARIANT TORUS THEOREM OF KOLMOGOROV

JACQUES FÉJOZ

ABSTRACT. The invariant torus theorem is proved using a simple fixed point theorem.

Let $\mathcal{H}$ be the space of germs along $\mathrm{T}_0^n := \mathbb{T}^n \times \{0\}$ of real analytic Hamiltonians in $\mathbb{T}^n \times \mathbb{R}^n = \{(\theta, r)\}$ ($\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$), endowed with the usual, inductive limit topology (see section 1). The vector field associated with $H \in \mathcal{H}$ is

$$\vec{H}: \quad \dot{\theta} = \partial_r H, \quad \dot{r} = -\partial_\theta H.$$

For $\alpha \in \mathbb{R}^n$, let $\mathcal{K}^\alpha$ be the affine subspace of Hamiltonians $K \in \mathcal{H}$ such that $K|_{\mathrm{T}_0^n}$ is constant (i.e. T_0^n is invariant) and $\vec{K}|_{\mathrm{T}_0^n} = \alpha$:

$$\mathcal{K}^\alpha = \{K \in \mathcal{H},\ \exists c \in \mathbb{R},\ K(\theta, r) = c + \alpha \cdot r + O(r^2)\}, \quad \alpha \cdot r = \alpha_1 r_1 + \cdots + \alpha_n r_n,$$

where $O(r^2)$ are terms of the second ordrer in r, which depend on θ.

Let also $\mathcal{G}$ be the space of germs along T_0^n of real analytic symplectomorphisms G in $\mathbb{T}^n \times \mathbb{R}^n$ of the following form:

$$G(\theta, r) = (\varphi(\theta), (r + \rho(\theta)) \cdot \varphi'(\theta)^{-1}),$$

where φ is an isomorphism of $\mathbb{T}^n$ fixing the origin (meant to straighten the flow on an invariant torus), and ρ is a closed 1-form on $\mathbb{T}^n$ (meant to straighten an invariant torus).

In the whole paper we fix $\alpha \in \mathbb{R}^n$ Diophantine ($0 < \gamma \ll 1 \ll \tau$; see [4]):

$$|k \cdot \alpha| \geq \gamma |k|^{-\tau} \quad (\forall k \in \mathbb{Z}^n \setminus \{0\}), \quad |k| = |k_1| + \cdots + |k_n|$$

and

$$K^o(\theta, r) = c^o + \alpha \cdot r + Q^o(\theta) \cdot r^2 + O(r^3) \in \mathcal{K}^\alpha$$

such that the average of the quadratic form valued function Q^o be non-degenerate:

$$\det \int_{\mathbb{T}^n} Q^o(\theta)\, d\theta \neq 0.$$

Theorem 1 (Kolmogorov [3, 1]). *For every $H \in \mathcal{H}$ close to K^o, there exists a unique $(K, G) \in \mathcal{K}^\alpha \times \mathcal{G}$ close to (K^o, id) such that $H = K \circ G$ in some neighborhood of $G^{-1}(\mathrm{T}_0^n)$.*

See [4, 5] and references therein for background. The functional setting below is related to [2].

1. The action of a group of symplectomorphisms

Define complex extensions $\mathbb{T}_\mathbb{C}^n = \mathbb{C}^n/\mathbb{Z}^n$ and $\mathrm{T}_\mathbb{C}^n = \mathbb{T}_\mathbb{C}^n \times \mathbb{C}^n$, and neighborhoods ($0 < s < 1$)

$$\mathbb{T}_s^n = \{\theta \in \mathbb{T}_\mathbb{C}^n,\ \max_{1 \leq j \leq n} |\mathrm{Im}\,\theta_j| \leq s\} \quad \text{and} \quad \mathrm{T}_s^n = \{(\theta, r) \in \mathrm{T}_\mathbb{C}^n,\ \max_{1 \leq j \leq n} \max(|\mathrm{Im}\,\theta_j|, |r_j|) \leq s\}.$$

For complex extensions U and V of real manifolds, denote by $\mathcal{A}(U, V)$ the Banach space of real holomorphic maps from the interior of U to V, which extend continuously on U; $\mathcal{A}(U) := \mathcal{A}(U, \mathbb{C})$.

• Let $\mathcal{H}_s = \mathcal{A}(\mathrm{T}_s^n)$ with norm $|H|_s := \sup_{(\theta,r) \in \mathrm{T}_s^n} |H(\theta, r)|$, such that $\mathcal{H} = \cup_s \mathcal{H}_s$ be their inductive limit.

Fix s_0. There exist ϵ_0 such that $K^o \in \mathcal{H}_{s_0}$ and, for all $H \in \mathcal{H}_{s_0}$ such that $|H - K^o|_{s_0} \le \epsilon_0$,

$$\left|\det \int_{\mathbb{T}^n} \frac{\partial^2 H}{\partial r^2}(\theta, 0)\, d\theta\right| \ge \frac{1}{2}\left|\det \int_{\mathbb{T}^n} \frac{\partial^2 K^o}{\partial r^2}(\theta, 0)\, d\theta\right| \ne 0. \tag{1}$$

Hereafter we assume that s is always $\ge s_0$. Set $\mathcal{K}^\alpha_s = \{K \in \mathcal{H}_s \cap \mathcal{K}^\alpha,\ |K - K^o|_{s_0} \le \epsilon_0\}$, and let $\vec{\mathcal{K}}_s \equiv \mathbb{R} \oplus O(r^2)$ be the vector space directing $\mathcal{K}^\alpha_s$.

• Let $\mathcal{D}_s$ be the space of isomorphisms $\varphi \in \mathcal{A}(\mathbb{T}^n_s, \mathbb{T}^n_{\mathbb{C}})$ with $\varphi(0) = 0$ and $\mathcal{Z}_s$ be the space of bounded real holomorphic closed 1-forms on $\mathbb{T}^n_s$. The semi-direct product $\mathcal{G}_s = \mathcal{Z}_s \rtimes \mathcal{D}_s$ acts faithfully and symplectically on the phase space by

$$G : \mathbb{T}^n_s \to \mathbb{T}^n_{\mathbb{C}}, \quad (\theta, r) \mapsto (\varphi(\theta), (\rho(\theta) + r) \cdot \varphi'(\theta)^{-1}), \quad G = (\rho, \varphi), \tag{2}$$

and, to the right, on $\mathcal{H}_s$ by $\mathcal{H}_s \to \mathcal{A}(G^{-1}(\mathbb{T}^n_s))$, $K \mapsto K \circ G$.

• Let $\mathfrak{d}_s := \{\dot\varphi \in \mathcal{A}(\mathbb{T}^n_s)^n,\ \dot\varphi(0) = 0\}$ with norm $|\dot\varphi|_s := \max_{\theta \in \mathbb{T}^n_s} \max_{1 \le j \le n} |\dot\varphi_j(\theta)|$, be the space of vector fields on $\mathbb{T}^n_s$ which vanish at 0. Similarly, let $|\dot\rho|_s = \max_{\theta \in \mathbb{T}^n_s} \max_{1 \le j \le n}, |\dot\rho_j(\theta)|$ on $\mathcal{Z}_s$. An element $\dot G = (\dot\rho, \dot\varphi)$ of the Lie algebra $\mathfrak{g}_s = \mathcal{Z}_s \oplus \mathfrak{d}_s$ (with norm $|(\dot\rho, \dot\varphi)|_s = \max(|\dot\rho|_s, |\dot\varphi|_s)$) identifies with the vector field

$$\dot G : \mathbb{T}^n_s \to \mathbb{C}^n, \quad (\theta, r) \mapsto (\dot\varphi(\theta), \dot\rho(\theta) - r \cdot \dot\varphi'(\theta)), \tag{3}$$

whose exponential is denoted by $\exp \dot G$. It acts infinitesimally on $\mathcal{H}_s$ by $\mathcal{H}_s \to \mathcal{H}_s$, $K \mapsto K' \cdot \dot G$.

Constants $\gamma_i, \tau_i, c_i, t_i$ below do not depend on s or σ.

Lemma 0. *If $\dot G \in \mathfrak{g}_{s+\sigma}$ and $|\dot G|_{s+\sigma} \le \gamma_0 \sigma^2$, then $\exp \dot G \in \mathcal{G}_s$ and $|\exp \dot G - \mathrm{id}\,|_s \le c_0 \sigma^{-1} |\dot G|_{s+\sigma}$.*

Proof. Let $\chi_s = \mathcal{A}(\mathbb{T}^n_s)^{2n}$, with norm $\|v\|_s = \max_{\theta \in \mathbb{T}^n_s} \max_{1 \le j \le n} |v_j(\theta)|$. Let $\dot G \in \mathfrak{g}_{s+\sigma}$ with $|\dot G|_{s+\sigma} \le \gamma_0 \sigma^2$, $\gamma_0 := (36n)^{-1}$. Using definition (3) and Cauchy's inequality, we see that if $\delta := \sigma/3$,

$$\|\dot G\|_{s+2\delta} = \max\left(|\dot\varphi|_{s+2\delta}, |\dot\rho + r \cdot \dot\varphi'(\theta)|_{s+2\delta}\right) \le 2n\delta^{-1} |\dot G|_{s+3\delta} \le \delta/2.$$

Let $D_s = \{t \in \mathbb{C},\ |t| \le s\}$ and $F := \left\{f \in \mathcal{A}(D_s \times \mathbb{T}^n_s)^{2n},\ \forall (t, \theta) \in D_s \times \mathbb{T}^n_s,\ |f(t, \theta)|_s \le \delta\right\}$. By Cauchy's inequality, the Lipschitz constant of the Picard operator

$$P : F \to F, \quad f \mapsto Pf, \quad (Pf)(t, \theta) = \int_0^t \dot G(\theta + f(s, \theta))\, ds$$

is $\le 1/2$. Hence, P possesses a unique fixed point $f \in F$, such that $f(1, \cdot) = \exp(\dot G) - \mathrm{id}$ and $|f(1, \cdot)|_s \le \|\dot G\|_{s+\delta} \le c_0 \sigma^{-1} |\dot G|_{s+\sigma}$, $c_0 = 6n$.

Also, $\exp \dot G \in \mathcal{G}_s$ because at all times the curve $\exp(t\dot G)$ is tangent to $\mathcal{G}_s$, locally a closed submanifold of $\mathcal{A}(\mathbb{T}^n_s, \mathbb{T}^n_{\mathbb{C}})$ (the method of the variation of constants gives an alternative proof). □

2. A property of infinitesimal transversality

We will show that locally $\vec{\mathcal{K}}_s$ is tranverse to the infinitesimal action of $\mathfrak{g}_s$ on $\mathcal{H}_{s+\sigma}$.

Lemma 1. *For all $(K, \dot H) \in \mathcal{K}^\alpha_{s+\sigma} \times \mathcal{H}_{s+\sigma}$, there exists a unique $(\dot K, \dot G) \in \vec{\mathcal{K}}_s \times \mathfrak{g}_s$ such that*

$$\dot K + K' \cdot \dot G = \dot H \quad \text{and} \quad \max(|\dot K|_s, |\dot G|_s) \le c_1 \sigma^{-t_1} (1 + |K|_{s+\sigma}) |\dot H|_{s+\sigma}.$$

Proof. We want to solve the linear equation $\dot K + K' \cdot \dot G = \dot H$. Write

$$\begin{cases} K(\theta, r) = c + \alpha \cdot r + Q(\theta) \cdot r^2 + O(r^3) \\ \dot K(\theta, r) = \dot c + \dot K_2(\theta, r), & \dot c \in \mathbb{R}, \quad \dot K_2 \in O(r^2) \\ \dot G(\theta, r) = (\dot\varphi(\theta), R + S'(\theta) - r \cdot \dot\varphi'(\theta)), & \dot\varphi \in \chi_s, \quad \dot R \in \mathbb{R}^n, \quad \dot S \in \mathcal{A}(\mathbb{T}^n_s). \end{cases}$$

Expanding the equation in powers of r yields

$$(4) \qquad \Big(\dot c + (\dot R + \dot S')\cdot\alpha\Big) + r\cdot\Big(-\dot\varphi'\cdot\alpha + 2Q\cdot(\dot R + \dot S')\Big) + \dot K_2 = \dot H =: \dot H_0 + \dot H_1\cdot r + O(r^2),$$

where the term $O(r^2)$ on the right hand side does not depend on $\dot K_2$.

Fourier series and Cauchy's inequality show that if $g \in \mathcal{A}(\mathbb{T}^n_{s+\sigma})$ has zero average, there is a unique function $f \in \mathcal{A}(\mathbb{T}^n_s)$ of zero average such that $L_\alpha f := f'\cdot\alpha = g$, and $|f|_s \le c\sigma^{-t}|g|_{s+\sigma}$ [4].

Equation (4) is triangular in the unknowns and successiveley yields:

$$\begin{cases} \dot S &= L_\alpha^{-1}\Big(\dot H_0 - \int_{\mathbb{T}^n}\dot H_0(\theta)\,d\theta\Big)\\ \dot R &= \frac12\left(\int_{\mathbb{T}^n}Q(\theta)\,d\theta\right)^{-1}\int_{\mathbb{T}^n}\Big(\dot H_1(\theta) - 2Q(\theta)\cdot\dot S'(\theta)\Big)\,d\theta\\ \dot\varphi &= L_\alpha^{-1}\Big(\dot H_1(\theta) - 2Q(\theta)\cdot(\dot R + \dot S'(\theta))\Big)\\ \dot c &= \int_{\mathbb{T}^n}\dot H_0(\theta)\,d\theta - \dot R\cdot\alpha\\ \dot K_2 &= O(r^2),\end{cases}$$

and, together with Cauchy's inequality, the wanted estimate. □

3. The local transversality property

Let us bound the discrepancy between the action of $\exp(-\dot G)$ and the infinitesimal action of $-\dot G$.

Lemma 2. *For all $(K,\dot H) \in \mathcal{K}^\alpha_{s+\sigma}\times\mathcal{H}_{s+\sigma}$ such that $(1+|K|_{s+\sigma})|\dot H|_{s+\sigma} \le \gamma_2\sigma^{\tau_2}$, if $(\dot K,\dot G) \in \vec{\mathcal{K}}\times\mathfrak{g}_s$ solves the equation $\dot K + K'\circ\dot G = \dot H$ (lemma 1), then $\exp\dot G \in \mathcal{G}_s$, $|\exp\dot G - \mathrm{id}\,|_s \le \sigma$ and*

$$|(K+\dot H)\circ\exp(-\dot G) - (K+\dot K)|_s \le c_2\sigma^{-t_2}(1+|K|_{s+\sigma})^2|\dot H|^2_{s+\sigma}.$$

Proof. Set $\delta = \sigma/2$. Lemmas 0 and 1 show that, under the hypotheses for some constant γ_2 and for $\tau_2 = t_1 + 1$, we have $|\dot G|_{s+\delta} \le \gamma_0\delta^2$ and $|\exp\dot G - \mathrm{id}\,|_s \le \delta$.

Let $H = K + \dot H$. Taylor's formula says

$$\mathcal{H}_s \ni H\circ\exp(-\dot G) = H - H'\cdot\dot G + \left(\int_0^1 (1-t)\,H''\circ\exp(-t\dot G)\,dt\right)\cdot\dot G^2$$

or, using the fact that $H = K + \dot K + K'\cdot\dot G$,

$$H\circ\exp(-\dot G) - (K+\dot K) = -(\dot K + K'\cdot\dot G)'\cdot\dot G + \left(\int_0^1 (1-t)\,H''\circ\exp(-t\dot G)\,dt\right)\cdot\dot G^2.$$

The wanted estimate thus follows from the estimate of lemma 1 and Cauchy's inequality. □

Let $B_{s,\sigma} = \{(K,\dot H) \in \mathcal{K}^\alpha_{s+\alpha}\times\mathcal{H}_{s+\sigma},\ |K|_{s+\sigma} \le \epsilon_0,\ |\dot H|_{s+\sigma} \le (1+\epsilon_0)^{-1}\gamma_2\sigma^{\tau_2}\}$ (recall (1)).

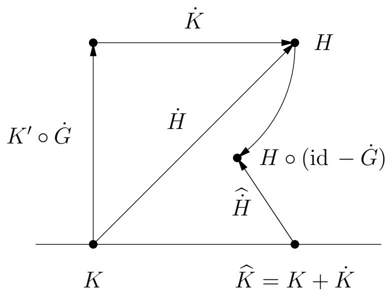
According to lemmas 1-2, the map $\phi : B_{s,\sigma} \to \mathcal{K}^\alpha_s\times\mathcal{H}_s$,

$$\phi(K,\dot H) = (K+\dot K, (K+\dot H)\circ\exp(-\dot G) - (K+\dot K)),$$

satisfies, if $(\hat K,\widehat{\dot H}) = \phi(K,\dot H)$,

$$|\hat K - K|_s, |\widehat{\dot H}|_s \le c_3\sigma^{-t_3}|\dot H|^2_{s+\sigma}.$$

Theorem 2 applies and shows that if $H - K^o$ is small enough in $\mathcal{H}_{s+\sigma}$, the sequence $(K_j,\dot H_j) = \phi^j(K^o, H - K^o)$, $j \ge 0$, converges towards some $(K,0)$ in $\mathcal{K}^\alpha_s\times\mathcal{H}_s$.



Let us keep track of the $\dot G_j$'s solving with the $\dot K_j$'s the successive linear equations $\dot K_j + K_j'\cdot\dot G_j = \dot H_j$ (lemma 1). At the limit,

$$K := K^o + \dot K_0 + \dot K_1 + \cdots = H\circ\exp(-\dot G_0)\circ\exp(-\dot G_1)\circ\cdots.$$

Moreover, lemma 1 shows that $|\dot{G}_j|_{s_{j+1}} \leq c_4 \sigma_j^{-t_4} |\dot{H}_j|_{s_j}$, hence the isomorphisms $\gamma_j := \exp(-\dot{G}_0)\circ \cdots \circ \exp(-\dot{G}_j)$, which satisfy

$$|\gamma_n - \mathrm{id}\,|_{s_{n+1}} \leq |\dot{G}_0|_{s_1} + \ldots + |\dot{G}_n|_{s_{n+1}},$$

form a Cauchy sequence and have a limit $\gamma \in \mathcal{G}_s$. At the expense of decreasing $|H - K^o|_{s+\sigma}$, by the inverse function theorem, $G := \gamma^{-1}$ exists in $\mathcal{G}_{s-\delta}$ for some $0 < \delta < s$, so that $H = K \circ G$.

Appendix. A fixed point theorem

Let $(E_s, |\cdot|_s)_{0<s<1}$ and $(F_s, |\cdot|_s)_{0<s<1}$ be two decreasing families of Banach spaces with increasing norms. On $E_s \times F_s$, set $|(x, y)|_s = \max(|x|_s, |y|_s)$. Fix $C, \gamma, \tau, c, t > 0$.

Let

$$\phi : B_{s,\sigma} := \{(x, y) \in E_{s+\sigma} \times F_{s+\sigma},\ |x|_{s+\sigma} \leq C,\ |y|_{s+\sigma} \leq \gamma\sigma^\tau\} \to E_s \times F_s$$

be a family of operators commuting with inclusions, such that if $(X, Y) = \phi(x, y)$,

$$|X - x|_s,\ |Y|_s \leq c\sigma^{-t}|y|^2_{s+\sigma}.$$

In the proof of theorem 1, "$|x|_{s+\sigma} \leq C$" allows us to bound the determinant of $\int_{\mathbb{T}^n} Q(\theta)d\theta$ away from 0, while "$|y|_{s+\sigma} \leq \gamma\sigma^\tau$" ensures that $\exp \dot{G}$ is well defined.

Theorem 2. *Given $s < s + \sigma$ and $(x, y) \in B_{s,\sigma}$ such that $|(x, y)|_{s+\sigma}$ is small, the sequence $(\phi^j(x, y))_{j\geq 0}$ exists and converges towards a fixed point $(\xi, 0)$ in $B_{s,0}$.*

Proof. It is convenient to first assume that the sequence is defined and $(x_j, y_j) := F^j(x, y) \in B_{s_j,\sigma_j}$, for $s_j := s + 2^{-j}\sigma$ and $\sigma_j := s_j - s_{j+1}$. We may assume $c \geq 2^{-t}$, so that $d_j := c\sigma_j^{-t} \geq 1$. By induction, and using the fact that $\sum 2^{-k} = \sum k2^{-k} = 2$,

$$|y_j|_{s_j} \leq d_{j-1}|y_{j-1}|^2_{s_{j-1}} \leq \cdots \leq |y|^{2^j}_{s+\sigma} \prod_{0\leq k\leq j-1} d_k^{2^{k+1}} \leq \Big(|y|_{s+\sigma} \prod_{k\geq 0} d_k^{2^{-k-1}}\Big)^{2^j} = \left(c4^t\sigma^{-t}|y|_{s+\sigma}\right)^{2^j}.$$

Given that $\sum_{n\geq 0} \mu^{2^n} \leq 2\mu$ if $2\mu \leq 1$, we now see by induction that if $|(x, y)|_{s+\sigma}$ is small enough, (x_j, y_j) exists in B_{s_j,σ_j} for all $j \geq 0$, y_j converges to 0 in F_s and the series $x_j = x_0 + \sum_{0\leq k\leq j-1}(x_{k+1} - x_k)$ converges normally towards some $\xi \in E_s$ with $|\xi|_s \leq C$. $\square$

Thank you to Ivar Ekeland, Éric Séré and Alain Chenciner for pointing out a mistake and for illuminating discussions.

References

[1] L. Chierchia. A. N. Kolmogorov's 1954 paper on nearly-integrable Hamiltonian systems. A comment on: "On conservation of conditionally periodic motions for a small change in Hamilton's function". *Regul. Chaotic Dyn.*, 13(2):130–139, 2008.

[2] J. Féjoz and M. Garay. Un théorème sur les actions de groupes de dimension infinie. *Comptes Rendus Mathematique*, 348(7-8):427 – 430, 2010.

[3] A. N. Kolmogorov. On the conservation of conditionally periodic motions for a small change in Hamilton's function. *Dokl. Akad. Nauk SSSR (N.S.)*, 98:527–530, 1954.

[4] J. Pöschel. A lecture on the classical KAM theorem. In *Smooth ergodic theory and its applications (Seattle, WA, 1999)*, volume 69 of *Proc. Sympos. Pure Math.*, pages 707–732. Amer. Math. Soc., Providence, RI, 2001.

[5] M. B. Sevryuk. The classical KAM theory at the dawn of the twenty-first century. *Mosc. Math. J.*, 3(3):1113–1144, 1201–1202, 2003. Dedicated to V. I. Arnold on the occasion of his 65th birthday.